Brain and Cognition 125 (2018) 149–156

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Telling a truth to deceive: Examining executive control and reward-related processes underlying interpersonal deception

Liyang Sai[a], Haiyan Wu[b], Xiaoqing Hu[c,*], Genyue Fu[a,*]

[a] Institute of Psychological Science, Zhejiang Key Laboratory for Research in Assessment of Cognitive Impairments, Center for Cognition and Brain Disorders, Hangzhou Normal University, Hangzhou, China
[b] Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Science, Beijing, China
[c] Department of Psychology, The State Key Laboratory of Brain and Cognitive Science, The University of Hong Kong, Hong Kong, China

ARTICLE INFO

Keywords:
ERP
Executive control
Outcome evaluation
Deception
Interpersonal deception

ABSTRACT

Does deception necessarily involve false statements that are incompatible with the truth? In some cases, people choose truthful statements in order to mislead others. This type of deception has been investigated less. The current study employed event-related brain potentials (ERPs) to investigate the neurocognitive processes when both truthful and false statements were used to deceive others. We focused our ERP analysis on two stages: a decision making stage during which participants decided whether to tell a false or a truthful statement, and an outcome evaluation stage during which participants evaluated whether their deception had succeeded or not. Results showed that in the decision making stage, intentions to deceive elicited larger N200s and smaller P300s than an honest control condition. During the outcome evaluation stage, success/failure feedback in the deception condition elicited larger Reward positivity (RewP) and feedback-P300 than feedback after honest responses. Importantly, whether participants chose to tell false or true statements, did not further modulate executive control or reward-related processes. Taken together, these results suggest that during interpersonal deception, having deceptive intentions engages executive control and reward-related processes regardless of the veracity of statements.

1. Introduction

To deceive others, people may spontaneously tell a falsified statement that is inconsistent with the truth. However, this strategy may not be optimal when potential recipients are already aware of the senders' deceptive intentions and therefore may not believe the senders' messages. In this scenario, the senders could strategically choose a truthful statement so that the recipient would take the truth as false. To date, the neurocognitive processes underlying such strategic deception involving truthful statements remain unclear. The present study employed an interpersonal deception game in which people deceived their opponents using both true and false statements. Compared with previous deception studies that compared truthful vs. false responses, we were able to use this manipulation to compare deceptive (regardless of the veracity of statements) with honest responses.

Previous studies have examined neurocognitive processes underlying both instructed and voluntary deception. In instructed deception, participants are instructed to lie and give false statements (e.g., deny their involvement of previous acts, see Abe et al., 2006; Lee et al., 2002; Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003; Spence et al., 2001). In voluntary deception, participants choose whether to make honest or dishonest decisions and they can over-report their performance for incentives (see Abe & Greene, 2014; Cui et al., 2018; Greene & Paxton, 2009; Hu, Pornpattananangkul, & Nusslock, 2015; Sip et al., 2010; Yin, Reuter, & Weber, 2016). Although instructed and voluntary deception differ along important dimensions such as social and motivational processes (e.g., perspective taking and reward processing, see Lisofsky, Kazzer, Heekeren, & Prehn, 2014), both deceptions require participants to execute a falsified response that is inconsistent with the truth. Specifically, the execution of truth-inconsistent responses requires the detection of conflict between two competing responses and then the inhibition of the goal-irrelevant truthful response (Abe et al., 2006; Greene & Paxton, 2009; Hu et al., 2015; Hu, Wu, & Fu, 2011; Johnson, Henkell, Simon, & Zhu, 2008; Johnson, Barnhardt, & Zhu, 2004). Based on these results, researchers hypothesize that truth-telling is the default response tendency (Christ, Van Essen, Watson, Brubaker, & McDermott, 2009; Farah, Hutchinson, Phelps & Wagner, 2014; Vrij, 2008; Vrij, Fisher, Mann, & Leal, 2006; but see Bereby-Meyer & Shalvi,

* Corresponding authors at: Department of Psychology, The University of Hong Kong, Pokfulam, Hong Kong, China (X. Hu). Institute of Psychological Science, Hangzhou Normal University, Hangzhou 311121, China (G. Fu).
E-mail addresses: xiaoqinghu@hku.hk (X. Hu), fugenyue@hznu.edu.cn (G. Fu).

https://doi.org/10.1016/j.bandc.2018.06.009
Received 11 January 2018; Received in revised form 26 June 2018; Accepted 27 June 2018

2015).

However, deception can be achieved by truthful statements as well. Critically, the defining feature of deception is that the message sender has an intention to mislead the recipient (Vrij, 2008). According to this definition, a deceiver could intentionally tell a truthful statement in order to mislead the recipient to believe its opposite. This type of deception can be adaptive especially when the recipient is already aware of the sender's deceptive intention, for example, in highly competitive scenarios such as negotiation (Rogers, Zeckhauser, Gino, Norton, & Schweitzer, 2017).

To date, very few studies have examined this type of deception (Carrión, Keenan, & Sebanz, 2010; Ding, Sai, Fu, Liu, & Lee, 2014; Sip et al., 2010; Volz, Vogeley, Tittgemeyer, von Cramon, & Sutter, 2015). Employing different methodologies such as ERP, fMRI, fNIRS, researchers consistently found that when the participants' goal was to mislead their opponents, telling a true statement would still engage similar executive control processes as telling a false statement. For example, in Sip et al. (2010), participants played the zero-sum dice game, Meyer, with a confederate. The participants' goal was to deceive the confederate about a dice combination. Sometimes participants chose to tell the truth about the dice combination, however, this was done with the expectation that the opponent would not trust them and thus believe the opposite to be true. Sip et al. (2010) found that both true and false claims about the dice combination were associated with higher activities in the fronto-polar cortex than that in a non-competitive control condition. Moreover, relative to truthful claims, false claims were associated with greater activity in the premotor and parietal cortices, which was taken as evidence that choosing a false claim additionally engaged response selection processes. Employing ERPs, Carrión et al. (2010) showed that both truthful and false claims with a deceptive intention elicited larger executive control-related ERPs, the medial frontal negativity (MFN), than truthful responses without deceptive intentions. These findings provide initial evidence that when truthful statements are used to deceive others, it involves similar executive control processes as when telling false statements to deceive.

Furthermore, because deception involves both information management (e.g. decision making) as well as risk management (e.g. outcome evaluation, see Sip, Roepstorff, McGregor, & Frith, 2008), the present study aims to extend previous research by examining both decision making and outcome evaluation processes in truth-telling deception. Critically, during interpersonal deception, a deceiver may not only decide whether or when to tell a false or a truthful statement, but the deceiver also needs to evaluate whether the deception has succeeded or not. This latter outcome evaluation stage may tap into reward-related processes (Hu et al., 2015; Luo, Sun, Mai, Gu, & Zhang, 2011; Sun, Chan, Hu, Wang, & Lee, 2015). To capture these two essential aspects of interpersonal deception, we leverage ERPs' high temporal resolution in a zero-sum, interpersonal deception game. Examining both stages of deception allows us to provide a more complete picture regarding interpersonal deception and its underlying neurocognitive mechanisms.

In the decision making stage, we focused on the fronto-central N200 and the centro-parietal P300, both of which are implicated in executive control processes. Specifically, it has been suggested that the N200 is a sensitive neural marker of response conflict (Bartholow et al., 2005; for a review, see Folstein & Van Petten, 2008); while the later P300 has been associated with cognitive resource allocation and conflict resolution (Johnson, 1988; Johnson, Barnhardt, & Zhu, 2003). Examining these two ERP components would also be consistent with previous ERP studies on both instructed and voluntary deception (Hu et al., 2015; Hu et al., 2011; Johnson et al., 2004; Johnson et al., 2008; Suchotzi, Crombez, Smulders, Meijer, & Verschuere, 2015; Wu, Hu, & Fu, 2009). Regarding how deception may modulate N200-P300, we hypothesize that telling a truth to deceive would engage similar executive control processes as when telling a lie to deceive (see Carrión et al., 2010). Moreover, deceptive responses, regardless of whether they were true or false, would elicit a larger N200 and smaller P300 than honest responses without deceptive intentions.

Regarding the outcome evaluation stage, we focused on two ERP components that have been intensively studied in the outcome evaluation literature: the Reward Positivity (RewP) and the feedback-P300. The RewP is typically observed during the 200–300 ms time window after the onset of the performance feedback, which indicates whether participants' behavior has led to good or bad outcomes (for a review, see Proudfit, 2015). Specifically, positive feedback would enhance this RewP while negative feedback would attenuate this RewP (Gehring & Willoughby, 2002; Miltner, Braun, & Coles, 1997; for reviews, see Proudfit, 2015; Walsh & Anderson, 2012).

The feedback-P300 is another ERP component that occurs later than RewP and is also intensively studied in the outcome evaluation research. Compared to the RewP, the results of feedback-P300 are less consistent across studies: Some studies have found that the feedback-P300 is sensitive to reward magnitude but not to valence (Sato et al., 2005; Yeung & Sanfey, 2004); while other studies have found that the feedback-P300 does encodes reward valence but may also implicate more cognitive processing of the feedback (Hajcak, Holroyd, Moser, & Simons, 2005; Hajcak, Moser, Holroyd, & Simons, 2007).

In relation to deception, Luo et al. (2011) employed an instructed deception paradigm and reported that the outcomes after instructed deception had elicited a larger RewP and feedback-P300 than the outcomes after honest responses (Luo et al., 2011). This finding suggests that instructed deception versus honesty modulates outcome evaluation processes. Based on this study, we also predicted that feedback after deception would elicit larger RewP and feedback-P300 than feedback after honest responses. Moreover, since both truthful and false statements serve the same goal to deceive others, we predict that there are no significant differences for RewP and feedback-P300 between truth-deceive and false-deceive responses.

2. Methods

2.1. Participants

Twenty-one undergraduates from Zhejiang Normal University were recruited in the study (9 males, Mean age = 24.05 years, age range 21–29 years). Two participants were excluded from behavioral and ERP analyses due to excessive eye blinks in the experiment; one additional participant was excluded from analyses in the outcome evaluation stage because he or she had insufficient outcome evaluation trials ($n < 20$) in the four conditions of the experiment. Thus, the final sample for the decision-making stage was 19 (7 males, Mean age = 24.16 years, age range 21–29 years), and the final sample for the outcome evaluation stage was 18 (7 males, Mean age = 24.17 years, age range 21–29 years). This sample size was consistent with previous studies on the topic ($n = 11$ in Carrión et al., 2010, $n = 25$ in Ding et al., 2014; $n = 14$ in Sip et al., 2010). All participants were right-handed with normal or corrected-to-normal vision. The study was approved by the Ethics Committee of Zhejiang Normal University.

2.2. Procedure

Participants completed two task sessions: an honest control session in which no deceptive intentions were involved; and an interpersonal deceptive game session during which participants were asked to mislead their opponents. These two sessions were presented in a fixed order, with the honest control session always coming first, followed by the interpersonal deception session. We chose this fixed order because if the honest session followed the interpersonal deception session, participants' honest behavior might be influenced by their previous deceptive intentions even when no deception was required (for a similar task order and rationale, see Carrión et al., 2010). In the interpersonal deceptive game, participants were told that they were about to play a